



日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

2000年 1月26日

出願番号

Application Number:

特願2000-016582

出願人

Applicant(s):

株式会社日立製作所

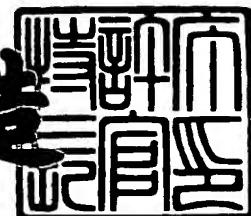
株式会社日立インフォメーションテクノロジー

Docket No.: ASA-959
Serial No.: 091767,707
Filed On: January 24, 2001
Mattingly, Stanger & Malin, P.C.
Phone: 703-684-1120

2001年 2月16日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3007463

【書類名】 特許願

【整理番号】 P0178JP

【提出日】 平成12年 1月26日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 12/46

【発明者】

【住所又は居所】 神奈川県秦野市堀山下 1 番地 株式会社日立製作所 エ
ンタープライズサーバ事業部内

【氏名】 須貝 和雄

【発明者】

【住所又は居所】 神奈川県秦野市堀山下 1 番地 株式会社日立インフォメ
ーションテクノロジー

【氏名】 松山 信仁

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【特許出願人】

【識別番号】 000153454

【氏名又は名称】 株式会社日立インフォメーションテクノロジー

【代理人】

【識別番号】 100107010

【弁理士】

【氏名又は名称】 橋爪 健

【手数料の表示】

【予納台帳番号】 054885

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ネットワーク中継装置

【特許請求の範囲】

【請求項 1】

ネットワーク内でパケットのヘッダ情報に従いパケットの中継先を決定しパケットを中継するネットワーク中継装置であり、
入力パケットの転送処理を行う複数のパケット転送ユニットと、
ルータのインターフェースから入力したパケットを各前記パケット転送ユニットに順番に振り分ける、又は、処理を行っていない前記パケット転送ユニットに振り分けるパケット振り分けユニットと、
各前記パケット転送ユニットがパケット転送処理を行ったパケットを、パケットの入力順に再配列するパケット再配列ユニット
を備えたネットワーク中継装置。

【請求項 2】

請求項 1 に記載のネットワーク中継装置において、
前記パケット振り分けユニットは、パケットのヘッダ条件からパケットを振り分ける前記パケット転送ユニットを決定することにより、前記パケット再配列ユニットでパケットの再配列をしない場合でもパケットの転送元から転送先に転送される一連のパケットの間でパケットの順序逆転が起きないようにすることを特徴とするネットワーク中継装置。

【請求項 3】

請求項 1 に記載のネットワーク中継装置において、
前記パケット振り分けユニットは、前記パケット転送ユニットにパケットを振り分ける時にパケットに通し番号を付け、
前記パケット転送ユニットは、パケット転送処理の間、パケットとともにパケットの通し番号を保持し、
前記パケット再配列ユニットは、パケットの通し番号の順にパケットを再配列することを特徴とするネットワーク中継装置。

【請求項 4】

請求項 1 に記載のネットワーク中継装置において、

前記パケット振り分けユニットがパケットを振り分けてからパケットの転送処理を開始するまでの間、処理を待つための第 1 のキューと、

前記パケット転送ユニットがパケット処理を終了してから前記パケット再配列ユニットがパケットを再配列するまでの間、処理を待つための第 2 のキューをさらに備えたネットワーク中継装置。

【請求項 5】

請求項 1 に記載のネットワーク中継装置において、

各前記パケット転送ユニット内で転送処理を行ったパケットのパケット数またはバイト数をカウントする第 1 のカウンタを備え、各前記パケット転送ユニット内で採取されたパケット数またはバイト数のカウント値を集計することにより、ネットワーク中継装置が転送処理を行ったパケット数またはバイト数を統計的に得ることを特徴とするネットワーク中継装置。

【請求項 6】

ネットワーク内でパケットのヘッダ情報に従いパケットの中継先を決定しパケットを中継するネットワーク中継装置であって、入力パケットの転送処理を行う複数個のパケット転送ユニットを備え、

前記パケット転送ユニットは、

パケットのヘッダの抽出および書き換えを行うパケットヘッダ操作機構と、

前記パケットヘッダ操作機構が抽出したパケットヘッダをキーとしてパケットヘッダの検索処理を行う複数個のパケット検索ユニットと、

複数の前記パケット検索ユニットにパケットヘッダを振り分ける検索パケット振り分け機構と、

複数の前記パケット検索ユニットが出力するパケットヘッダ検索結果を再配列する検索パケット再配列ユニット

を備えたネットワーク中継装置。

【請求項 7】

請求項 6 に記載のネットワーク中継装置において、

前記検索パケット振り分け機構は、

各前記パケット検索ユニットに順番にパケットヘッダを振り分ける、又は、処理を行っていない前記パケット検索ユニットにパケットヘッダを振り分けることを特徴とするネットワーク中継装置。

【請求項 8】

請求項 6 に記載のネットワーク中継装置において、

前記検索パケット振り分け機構は、

パケットのヘッダ条件からパケットを振り分ける前記パケット検索ユニットを決定することにより、前記検索パケット再配列ユニットでパケットの再配列をしない場合でもパケットの転送元から転送先に転送される一連のパケットの間でパケットの順序逆転が起きないようにすることを特徴とするネットワーク中継装置。

【請求項 9】

請求項 6 のネットワーク中継装置において、

前記検索パケット振り分けユニットは、前記パケット検索ユニットにパケットヘッダを振り分ける時にパケットヘッダに通し番号を付け、

前記パケット検索ユニットは、パケットヘッダ検索処理の間、パケットヘッダと一緒に該パケットヘッダの通し番号を保持し、

前記検索パケット再配列ユニットは、該パケットヘッダの通し番号の順に検索結果を再配列することを特徴とするネットワーク中継装置。

【請求項 1 0】

請求項 6 に記載のネットワーク中継装置において、

前記検索パケット振り分けユニットが、パケットヘッダを振り分けてからパケットヘッダの検索処理を開始するまでの間、処理を待つための第 3 のキューと、

前記パケット検索ユニットがパケットヘッダ検索処理を終了してから前記検索パケット再配列ユニットが検索結果を再配列するまでの間、処理を待つための第 4 のキューをさらに備えたネットワーク中継装置。

【請求項 1 1】

請求項 6 に記載のネットワーク中継装置において、

各前記パケット検索ユニット内で検索処理を行ったパケットのパケット数またはバイト数をカウントする第 2 のカウンタを備え、各前記パケット検索ユニット

内で採取されたパケット数またはバイト数のカウント値を集計することにより、ネットワーク中継装置が検索処理を行ったパケット数またはバイト数を統計的に得ることを特徴とするネットワーク中継装置。

【請求項 1 2】

請求項 6 に記載のネットワーク中継装置において、

前記パケット検索ユニットは、パケットのヘッダ条件に従いパケットを分類し、この各分類結果を特定する番号を検索結果として出力し、

前記検索パケット再配列ユニットは、前記パケット検索ユニットから入力されたパケットの各分類結果を特定する番号毎にパケットの転送速度を測定し、

この測定結果に従い、入力されたパケットが一定値以上だったらパケットを廃棄する処理、輻輳時に廃棄され易くする処理、または、出力するパケットの速度が一定値以上にならないようにパケットの出力速度を抑える処理等の処理を実行することを特徴とするネットワーク中継装置。

【請求項 1 3】

請求項 1 乃至 5 に記載のネットワーク中継装置において、請求項 6 乃至 1 3 のいずれかに記載のネットワーク中継装置における前記パケット転送ユニットを用いるようにしたネットワーク中継装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、ネットワーク中継装置に係る。本発明は、特に、コンピュータネットワークシステムにおけるルータ、リピータ等のネットワーク中継装置において、入力されたパケットの転送先を高速に検索することができるネットワーク中継装置に関する。

【0 0 0 2】

【従来の技術】

一般にネットワークシステムにおいては、複数のネットワーク間を接続するためにルータやブリッジ等のネットワーク中継装置が用いられる。ルータは、接続されているネットワークから受け取ったパケットの宛先アドレスを調べてパケット

の転送先を決定し、転送先のルータ又はホストが接続されたネットワークにパケットを転送する。

【0 0 0 3】

ルータは、ネットワークインターフェースボードからパケットを入力する毎にパケットのヘッダ情報に含まれるパケットの宛先情報からそのパケットをどのインターフェースに出力するかを決定し、さらに、出力するインターフェースに接続されるホスト又はルータの内、どのホストまたはルータに転送するかを決定して、パケットを転送する。さらに高機能のルータでは、パケットを転送するだけでなくパケットヘッダ条件に従いパケットの廃棄を行う機能(フィルタリング機能)、複数のパケットを転送するときパケットの出力優先度を決定し、パケットの優先度の順にパケットを出力する機能(優先制御機能)、パケットヘッダ内の優先度情報(T o S)を書き換える機能等を実行する。

【0 0 0 4】

さらに、一般にルータは統計機能を有しており、転送を行ったパケット数およびバイト数、ルータ内でフィルタリングやルータ内のパケットの輻輳により廃棄したパケット数、パケットの出力先の検索(経路検索)の結果、どこに出力するかが定義されていなかったパケットの数(経路無し)、ルータ自身に宛てたパケットの数や、フィルタリングの廃棄条件に一致したパケット数等を数えている。これらの統計情報は、ルータの操作や状態表示等の管理を行う端末に表示したり、ネットワーク内に存在するネットワーク管理装置に送られ、ネットワークの管理に使用される。

【0 0 0 5】

また、統計機能ではないが、パケットのデータ転送バイト数の測定が必要な場合として、ネットワーク業者がユーザからのパケットを受信する時にパケットのヘッダ条件に従いパケットを分類し(例えば、ソース・デスティネーションのサブネット、TCP/IPのポート番号、アドレス等)、この分類(以後フローと呼ぶ)毎にデータ転送速度の監視を行い、このフローのデータ転送速度がこのフローを流しているユーザとの間で契約している帯域以上のパケットを廃棄したり輻輳時に優先して廃棄するようにする機能(帯域監視機能)、または、ネットワーク

ユーザがネットワーク業者のネットワークにパケットを送信する時にルータから各フローに属するパケットデータを出力する速度をネットワーク業者との間で契約している帯域以内に制限する機能(シェーピング機能)がある。帯域監視機能またはシェーピング機能を実現するためには、ルータを流れるパケットの一定時間毎のデータ転送バイト数(データ転送速度)を測定する必要がある。

【0 0 0 6】

ルータの処理内容には、ルータにパケットが入ってくる毎に行う必要がある上述のような処理と、ルータ間でネットワークシステム内の経路情報をやりとりしルータ内に経路テーブルを作成するという処理や、ネットワーク管理装置との間でのルータの統計情報のやり取りのようにルータにパケットが入ってくる毎には行う必要がない処理がある。

【0 0 0 7】

従来、上記ルータの処理はソフトウェア処理により行われており、ルータにパケットが入ってくる毎に必要な処理の性能向上のため、この処理を行うユニット(RP: Routing Processor)をルータ内に複数設け、各RP毎にネットワークインターフェースボードを接続し、別々のRPの配下に繋がるネットワークインターフェースボード間でパケットを転送するためにRP間をバスまたはクロスバースイッチにより接続することにより高速化が図られていた。このようなルータの例として、例えば、池田、他の特開平5-199230号公報「インタネットワーク装置及び通信ネットワークシステム」等が挙げられる。

【0 0 0 8】

しかし、さらにネットワークを流れるトラフィックが増大しルータ処理を高速化する必要性が出てきた結果、ソフトウェア処理により実現されていたルータ処理をハードウェア処理で置き換えるということが行われるようになってきた。このようなルータの例として、日経エレクトロニクス1999年11月2日号pp. 35-36に示す日立製作所のルータ「GR2000」等が挙げられる。

【0 0 0 9】

【発明が解決しようとする課題】

近年のネットワークを流れるトラフィックは、LSI製造技術等のハードウェア

の要素技術の進歩により実現できる処理能力の向上以上のペースで増大し続けており、ルータやブリッジがネットワークインターフェースボードからパケットを入力する毎に行う処理をハードウェア化するだけでは増大するネットワークトラフィックを処理しきれない。

【0 0 1 0】

本発明は、以上の点に鑑み、ルータの処理を行うハードウェアを並列化することにより、L S I 製造技術等のハードウェアの要素技術の進歩により実現できる処理能力の向上以上にルータのトータル性能を上げ、近年の増大し続けるネットワークトラフィックを十分に処理することができるルータ装置を提供することを目的とする。

【0 0 1 1】

【課題を解決するための手段】

本発明の特徴のひとつとしては、ルータ内にパケットの入出力のインターフェースを行うネットワークインターフェースボード(N I F : Network Interface Board)、パケットのヘッダ情報からパケットの転送先を検索しパケットを転送するルーティングプロセッサ(R P : Routing Processor)、各 R P の制御およびルーティングプロトコル等の制御パケットの送受信を行うルーティングマネージャ(R M : Routing Manager)、および、各 R P および R M を接続し R P 間および R M、R P 間のパケットの転送を行うクロスバスイッチ(C S W : Cross bar Switch)を設け、パケットのヘッダ情報からパケットの宛先を検索しパケットを転送するという処理を実行するパケット転送ユニットを各 R P 内に複数個設け、N I F から入ってきたパケットおよび C S W から入ってきたパケットを、パケット単位にパケット転送ユニットに振り分けるようにする。

【0 0 1 2】

パケット転送ユニットが行う処理を限定しパケット転送ユニットの構成を簡単化するためには、N I F から入ってきたパケットを C S W に転送する処理を行うパケット転送ユニットと、C S W 経由で入ってきたパケットを N I F に転送する処理を行うパケット転送ユニットを別々に設けるようにすればよい。また、パケット転送ユニットが処理を行わないアイドル時間を短くしパケット転送ユニット当

りのルータの処理性能を向上させるためには、各パケット転送ユニットにN I FからC S Wへの転送とC S WからN I Fへの転送のどちらの転送処理も行わせるようにすればよい。

【 0 0 1 3 】

本発明の他の特徴のひとつとしては、各R PがN I FまたはC S Wから入ってきたパケットを各R P内の複数のパケット転送ユニットに振り分けるために、転送パケット振り分けユニットを設け、この転送パケット振り分けユニットは、パケットが入ってきた順にパケット転送ユニットに順番に振り分ける方法(ラウンドロビン法)か、空いているパケット転送ユニットに振り分ける方法(アウトオブオーダー法)か、パケットのヘッダ条件から振り分けるパケット転送ユニットを一意に決定する方法(ハッシュ法)によりパケットを振り分けるようにする。

【 0 0 1 4 】

ハッシュ法でパケットの転送先を決定する関数(ハッシュ関数)には、パケットの転送元から転送先に連続して送られる一連のパケットは全て同じパケット転送ユニットに振り分ける、という条件を設けることができる。さらに、各パケット転送ユニットにパケットができるだけ均等に振り分けられるように、ハッシュ関数としてパケットヘッダ条件から作成された擬似乱数を使用するようにすることができる。

【 0 0 1 5 】

また、本発明の他の特徴のひとつとして、各R P内に、R P内の複数のパケット転送ユニットが転送処理を行ったパケットを再び配列してN I FまたはC S Wに送出する転送パケット再配列ユニットを設ける。

【 0 0 1 6 】

転送パケット振り分けユニットの振り分け法がラウンドロビン法またはアウトオブオーダー法の場合には、R P内の複数のパケット転送ユニットがパケット転送処理を終了する順番が、パケットの転送元から転送先に連続して送られる一連のパケットの間で逆転することが有り得るので、転送パケット再配列ユニットは、各パケット転送ユニットがパケット転送処理を終了したパケットをパケットが入ってきた順に並べ直す。転送パケット振り分けユニットの振り分け法がアウトオブ

オーダ法の場合には、転送パケット再配列ユニットがルータにパケットが入ってきた順番にパケットを再配列するために、転送パケット振り分けユニットがパケット転送ユニットにパケットを振り分ける時点で各パケットにシーケンス番号を付け、シーケンス番号とパケットとを一緒にパケット転送ユニットの中を持ち回り、転送パケット再配列ユニットはパケットに付けられたシーケンス番号の順にパケットを再配列してパケットを送出するようにする。

【0 0 1 7】

また、本発明のさらに他の特徴のひとつとしては各パケット転送ユニット内には、転送パケット振り分けユニットからパケットを受けてからパケット転送処理を開始するまでの間待つためのキューである転送パケット処理待ちキューと、パケット転送処理ユニットが処理を終了してから転送パケット再配列ユニットにパケットを送出するまでの間待つためのキューである転送パケット再配列処理待ちキューを設けるようにする。

【0 0 1 8】

パケット転送ユニットが行う処理の内、パケットヘッダの検索を並列化して高速化を図るためには、一つのパケット転送ユニット内にパケットヘッダの検索処理を行うパケット検索ユニットを複数個設けるようにすることができる。また、パケット転送ユニット内でパケットのヘッダの抽出を行うパケットヘッダ操作機構がパケットのヘッダ情報を抽出するようにし、抽出したパケットヘッダを各パケット検索ユニットに振り分ける検索パケット振り分けユニットを設け、各パケット検索ユニットはパケットのヘッダ情報に基づきパケットの転送先、パケットを廃棄するか転送するかの検索(フィルタリング検索)、および、パケットを出力するときの出力順の優先順位(通信品質検索)を行い、パケットの転送先、パケットを廃棄するか否かの情報、パケットの出力優先順位を返すようにし、各パケット検索ユニットの検索結果を再び配列してパケットヘッダ操作機構に渡す検索パケット再配列ユニットを設けるようにすることができる。

【0 0 1 9】

さらに高速化を図るためには、パケット転送ユニットの並列化とパケット検索ユニットの並列化を同時に実行することもできる。

【 0 0 2 0 】

さらに、本発明の他の特徴のひとつとしては、複数のパケット転送ユニットまたは複数のパケット検索ユニットにそれぞれ採取された統計情報を集計しルータ全体の統計情報として管理端末に表示する、あるいはネットワーク管理装置に送るために、ルータ全体を管理するソフトウェアが各パケット転送ユニットまたはパケット検索ユニット内の統計情報を読み、ソフトウェア処理により集計することができる。あるいは、各 R P 毎にプロセッサを設け、このプロセッサ上で動作するソフトウェアが各パケット転送ユニットまたはパケット検索ユニット内の統計情報を読み、ソフトウェア処理により集計することができる。

【 0 0 2 1 】

また、本発明のさらに他の特徴のひとつとしては、パケットのヘッダ条件毎にパケットを分類し、この分類(以後フローと呼ぶ)毎にデータ転送速度を測定し、データ転送速度に従いパケットを廃棄したり輻輳時にパケットを廃棄され易くするため、あるいはパケットの送出時にパケットの送出速度を制限するために、フローに番号を付け、各データ転送ユニットまたは各データ検索ユニットからフローの番号を出力し、転送パケット再配列ユニットまたは検索パケット再配列ユニット内で、フローの番号毎にフローのデータ転送速度を測定することができる。

【 0 0 2 2 】

【発明の実施の形態】

図 1 に、本発明のルータ装置が適用されるネットワークシステム全体の概念図を示す。図 1 において、パケットの送信元および受信先になり得るホスト 1、2 および 3、パケットのヘッダ内に記述されている宛先アドレスに従いパケットを転送するルータ 4 および 5、ホストおよびルータが直接接続され、ホストおよびルータ間のデータを伝送するサブネットワーク 6、7 および 8 が示される。ルータ 4 および 5 もパケットの送信元および受信先になり得る。

【 0 0 2 3 】

例えば、ホスト 1 からホスト 2 にパケットを転送する場合、ホスト 1 はインター

フェース 1 0 からサブネットワーク 6 を介して繋がるルータ 4 に向かってパケットを送出する。ルータ 4 はサブネットワーク 6 に繋がるインターフェース 4 0 からパケットを受け取り、インターフェース 4 1 からサブネットワーク 7 を介して繋がるルータ 5 に向かってパケットを転送する。同様にルータ 5 はインターフェース 5 0 からパケットを受け取り、インターフェース 5 1 からサブネットワーク 8 を介して繋がるホスト 2 に向かってパケットを転送する。ホスト 2 はインターフェース 2 0 からパケットを受け取る。サブネットワーク 内の特定のルータまたはホストの特定のためにはサブネットワーク内のアドレスが使用される。

【 0 0 2 4 】

図 2 に、ルータ内部の全体ブロック図を示す。図 2 においてルータ 4 は、パケットの入出力のインターフェースを行うネットワークインターフェースボード(N I F : Network Interface Board) 4 5 0 ~ 4 5 5、パケットのヘッダ情報からパケットの転送先を検索しパケットを転送するルーティングプロセッサ(R P : Routing Processor) 4 4 0 ~ 4 4 2、ルータ全体の管理、各 R P の制御、および、ルーティングプロトコル等の制御パケットの送受信を行い経路情報の管理等を行うルーティングマネージャ(R M : Routing Manager) 4 2、および、各 R P および R M を接続し、R P 間および R M、R P 間のパケットの転送を行うクロスバスイッチ(C S W : Cross bar Switch) 4 3 を備える。各ネットワークインターフェースボード (N I F) 4 5 0 ~ 4 5 5 からはネットワークインターフェース 4 0 0 ~ 4 0 5 が繋がる。図 2 のネットワークインターフェース 4 0 0 ~ 4 0 5 は、図 1 のネットワークインターフェース 4 0、4 1 に対応するものである。ネットワークインターフェースの種類には、イーサネット等の L A N (Local Area Network)、W A N (Wide Area Network)、A T M (Asynchronous Transfer Mode) 等があり、各 N I F 4 5 0 ~ 4 5 5 はインターフェースの種類毎に異なるハードウェアで構成される。

【 0 0 2 5 】

図 3 および図 4 に R P 4 4 0 の内部のブロック図を示す。図 3 および図 4 は、パケット転送ユニットを並列化した本発明のルーティングプロセッサ内部の構成の 2 つの例を表す構成図である。パケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図

3) および 4 4 1 0～4 4 1 3 (図 4) は、R P 内でパケットのヘッダ情報からパケットの宛先を検索し、パケットを転送するという処理を実行するユニットであり、R P 内には負荷分散のために複数のパケット転送ユニットを設け、N I F から入ってきたパケットおよび C S W から入ってきたパケットをパケット単位にパケット転送ユニットに振り分けるようにする。図 3 は N I F から入ってきたパケットを転送するパケット転送ユニット 4 4 0 0～4 4 0 3 と C S W から入ってきたパケットを転送するパケット転送ユニット 4 4 0 4～4 4 0 7 を別々に設ける場合の R P の構成例であり、図 4 は各パケット転送ユニット 4 4 1 0～4 4 1 3 が N I F から入ってきたパケットと C S W から入ってきたパケットのどちらの転送処理も行う場合の R P の構成例である。

【0 0 2 6】

転送パケット振り分けユニット 4 4 2 0、4 4 2 2 (図 3) および 4 4 3 0、4 4 3 2 (図 4) は、N I F または C S W から入ってきたパケットを複数のパケット転送ユニットのいずれかに振り分けるという処理を行うユニットであり、パケットの振り分け方法として、パケットが入ってきた順にパケット転送ユニットに順番に振り分ける方法(ラウンドロビン法)と、空いているパケット転送ユニットに振り分ける方法(アウトオブオーダー法)と、パケットのヘッダ条件から振り分けるパケット転送ユニットを一意に決定する方法(ハッシュ法)等がある。図 1 1、図 1 2、図 1 3 にそれぞれ、転送パケット振り分けユニットのパケット振り分け方法が、ラウンドロビン法、アウトオブオーダー法、ハッシュ法の場合の各パケット転送ユニットのパケット転送処理を表すタイムチャートを示す。図 1 1、1 2、1 3 のタイムチャート内の番号は R P へのパケットの入力順を示し、各矩形は各パケット転送ユニットが各パケットを処理している時間を示す。図 1 1、1 2、1 3 では、処理時間が長く掛かるパケットが奇数番目に入ってきて、処理時間が短いパケットが偶数番目に入ってきた場合の例について示す。

【0 0 2 7】

ハッシュ法でパケットの転送先を決定する関数(ハッシュ関数)には、例えば、パケットの転送元から転送先に連続して送られる一連のパケットは全て同じパケット転送ユニットに振り分ける、という条件を設ける。この条件により、パケット

毎に転送処理に掛かる時間が異なった場合にパケット転送処理が終了した順にパケットを送出しても、パケットの転送元から転送先に連続して送られる一連のパケットの間で転送元からパケットを送った順に転送先に到着させることができる。

【 0 0 2 8 】

さらに、各パケット転送ユニットにパケットができるだけ均等に振り分けられるように、ハッシュ関数としてパケットヘッダ条件から作成された擬似乱数を使用することもできる。パケット転送ユニットの数が2個の場合のハッシュ関数の例として、32ビットのソースIPアドレスと32ビットのデスティネーションIPアドレスの全ビットの排他的論理和を取る、という方法がある。この場合、全ビットの排他的論理和により得られる1ビットにより2個の転送ユニットの内どちらに振り分けるかを決定する。パケット転送ユニットの数が4個の場合には1ビットでなく2ビットを出力するハッシュ関数を用いる。図13では1、3番目のパケットがパケット検索ユニット4410に振り分けられ、2、4番目が4411に、5、7番目が4412に、6、8番目が4413にそれぞれ振り分けられた場合の例について示している。

【 0 0 2 9 】

転送パケット振り分けユニット4420、4422（図3）および4430、4432（図4）の振り分け法がラウンドロビン法またはアウトオブオーダー法の場合には、RP内の複数のパケット転送ユニット4400～4407（図3）および4410～4413（図4）がパケット転送処理を終了する順番が、パケットの転送元から転送先に連続して送られる一連のパケットの間で逆転することが有り得るので、転送パケット再配列ユニット4421、4423（図3）および4431、4433（図4）は、各パケット転送ユニット4400～4407（図3）および4410～4413（図4）がパケット転送処理を終了したパケットをパケットが入ってきた順に並べ直す。また、転送パケット振り分けユニット4420、4422（図3）および4430、4432（図4）の振り分け法がアウトオブオーダー法の場合には、転送パケット再配列ユニット4421、4423（図3）および4431、4433（図4）がルータにパケットが入ってきた順

番にパケットを再配列するために、転送パケット振り分けユニット 4 4 2 0、4 4 2 2 (図 3) および 4 4 3 0、4 4 3 2 (図 4) がパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3) および 4 4 1 0 ~ 4 4 1 3 (図 4) にパケットを振り分ける時点で各パケットにシーケンス番号を付け、シーケンス番号とパケットとを一緒にパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3) および 4 4 1 0 ~ 4 4 1 3 (図 4) の中を持ち回り、転送パケット再配列ユニット 4 4 2 1、4 4 2 3 (図 3) および 4 4 3 1、4 4 3 3 (図 4) はパケットに付けられたシーケンス番号の順にパケットを再配列してパケットを送出する。

【0 0 3 0】

また、転送パケット振り分けユニット 4 4 2 0、4 4 2 2 (図 3) および 4 4 3 0、4 4 3 2 (図 4) の振り分け法がハッシュ法の場合には、各パケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3) および 4 4 1 0 ~ 4 4 1 3 (図 4) に送られたパケット間でパケットの出力順は前後して良く、転送パケット再配列ユニット 4 4 2 1、4 4 2 3 (図 3) および 4 4 3 1、4 4 3 3 (図 4) はパケット転送処理の終了順にパケットを送出する。複数のパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3) および 4 4 1 0 ~ 4 4 1 3 (図 4) のパケット転送処理が終了していた場合には、パケット転送処理が終了した複数のパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3) および 4 4 1 0 ~ 4 4 1 3 (図 4) 間で順番にパケットを送出する。

【0 0 3 1】

つぎに、図 5 に、本発明のパケット転送ユニットの内部構成図を示す。パケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3)、4 4 1 0 ~ 4 4 1 3 (図 4) は、パケットの転送方向、即ち N I F 4 5 0 から C S W 4 3 向きか、C S W 4 3 から N I F 4 5 0 向きかが異なるが内部の処理待ちのためのキューの持ち方は同じであり、図 5 は、代表してパケット転送ユニット 4 4 0 0 の内部の処理待ちのためのキューの持ち方を示すブロック図を表したものである。図 5 のパケット転送ユニット 4 4 0 0 内にはパケット転送処理を実行するパケット転送処理機構 4 4 0 0 0、転送パケット振り分けユニット 4 4 2 0 等が各パケット転送ユニット 4 4 0 0 にパケットを振り分けた後、各パケット転送処理機構 4 4 0 0 0 がパ

ケット処理を開始するまで処理を待つためのキューである転送パケット処理待ちキュー４４００１と、各パケット転送処理機構４４０００が処理を終了後、転送パケット再配列ユニット４４２１等がパケットを再配列し出力するまで待つためのキューである転送パケット再配列処理待ちキュー４４００２を備える。

【００３２】

転送パケット処理待ちキュー４４００１および転送パケット再配列処理待ちキュー４４００２は、各パケット転送ユニット４４００等の各パケットの処理時間の長短、および、転送パケット振り分けユニット４４２０等の振り分け方がハッシュ法の場合に各パケット転送ユニット４４００等にパケットを振り分けるときの時間的偏りを吸収し、各パケット転送ユニット４４００等が処理をしていない時間を少なくするという効果がある。

【００３３】

次に、パケット転送ユニット４４００内でパケットの検索処理を行う部分だけを並列化して高速化を図る場合の実施の形態を示す。まず、図６に、パケットヘッダの検索処理を行うパケット検索ユニットを並列化しない場合のパケット転送ユニット４４００およびその内部のパケット転送処理機構４４０００の内部構成図を示す。また、図７に、パケット検索ユニットを並列化した場合のパケット転送ユニット４４００およびその内部のパケット転送処理機構４４０００の内部構成図を示す。

【００３４】

図６のパケット転送処理機構４４０００は、パケットヘッダ操作機構４４００００とパケット検索ユニット４４０００１を備える。パケットヘッダ操作機構４４００００は、パケットのヘッダ部分を抽出し、パケット検索ユニット４４０００１に渡す。パケット検索ユニット４４０００１はパケットヘッダ操作機構４４００００から渡されたパケットヘッダ情報に基づきパケットの転送先、パケットを廃棄するか転送するかの検索（フィルタリング検索）、および、パケットを出力するときの出力順の優先順位（通信品質検索）等を行い、パケットの転送先、パケットを廃棄するか否かの情報、パケットの出力優先順位等の情報をパケットヘッダ操作機構４４００００に渡す。パケットヘッダ操作機構４４００００はパケ

ット検索ユニット4 4 0 0 0 1から渡されたパケットの転送先、パケットを廃棄するか転送するか等の情報に従い、パケットを転送、または廃棄する。

【0 0 3 5】

図7に、パケット検索ユニットを並列化した場合の、パケット転送ユニット4 4 0 0の構成を示す。このパケット転送ユニット4 4 0 0は、パケット転送処理機構4 4 0 0 0、パケットヘッダ操作機構4 4 0 0 0 0、図6のパケット検索ユニット4 4 0 0 0 1に対応するパケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4、パケット転送処理機構4 4 0 0 0が複数のパケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4の一つにパケットヘッダの検索処理を振り分ける検索パケット振り分けユニット4 4 0 0 0 5、各パケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4の検索結果を再配列する検索パケット再配列ユニット4 4 0 0 0 6を備える。

【0 0 3 6】

検索パケット振り分けユニット4 4 0 0 0 5のパケット検索ユニットの振り分け方法には、転送パケット振り分けユニット4 4 2 0、4 4 2 2（図3）、および4 4 3 0、4 4 3 2（図4）と同様、ラウンドロビン法とアウトオブオーダー法とハッシュ法を取り得て、それぞれの振り分け法の特徴も転送パケット振り分けユニットの場合と同様である。検索パケット再配列ユニット4 4 0 0 0 6の再配列方法も、転送パケット再配列ユニット4 4 2 1、4 4 2 3（図3）および4 4 3 1、4 4 3 3（図4）の再配列方法と同様である。

【0 0 3 7】

ここで、図8に、本発明のパケット検索ユニットの内部構成図を示す。

図7のパケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4は内部の処理待ちのためのキューの持ち方は同じであり、図8は、代表してパケット検索ユニット4 4 0 0 0 1の内部の処理待ちのためのキューの持ち方を示すブロック図を表したものである。図8のパケット検索ユニット4 4 0 0 0 1内にはパケット検索処理を実行するパケット検索処理機構4 4 0 0 0 1 0、検索パケット振り分けユニット4 4 0 0 0 5が各パケット検索ユニット4 4 0 0 0 1等にパケットを振り分けた後、各パケット検索処理機構4 4 0 0 0 1 0がパケット処理を開始するまで処理

を待つためのキューである検索パケット処理待ちキュー 4 4 0 0 0 1 1 と、各パケット検索処理機構が処理を終了後、検索パケット再配列ユニットがパケットを再配列し出力するまで待つためのキューである検索パケット再配列処理待ちキュー 4 4 0 0 0 1 2 を備える。

【 0 0 3 8 】

検索パケット処理待ちキュー 4 4 0 0 0 1 1 および検索パケット再配列処理待ちキュー 4 4 0 0 0 1 2 は、各パケット検索ユニット 4 4 0 0 0 1 等の各パケットの処理時間の長短、および、検索パケット振り分けユニットの振り分け方がハッシュ法の場合に各パケット検索ユニットにパケットを振り分けるときの時間的偏りを吸収し、各パケット検索ユニットが処理をしていない時間を少なくするという効果がある。

【 0 0 3 9 】

次に、上記パケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3)、4 4 1 0 ~ 4 4 1 3 (図 4) またはパケット検索ユニット 4 4 0 0 0 1 ~ 4 4 0 0 0 4 を並列化した場合の統計情報採取法について示す。例えばパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3)、4 4 1 0 ~ 4 4 1 3 (図 4) 内では転送を行ったパケット数およびバイト数、ルータ内でフィルタリングやルータ内のパケットの輻輳により廃棄したパケット数を数えており、また、パケット検索ユニット 4 4 0 0 0 1 ~ 4 4 0 0 0 4 内ではパケットの出力先の検索(経路検索)の結果、どこに出力するかが定義されていなかったパケットの数(経路無し)、ルータ自身に宛てたパケットの数や、フィルタリングの廃棄条件に一致したパケット数等を数えている。

【 0 0 4 0 】

本発明ではパケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3)、4 4 1 0 ~ 4 4 1 3 (図 4) またはパケット検索ユニット 4 4 0 0 0 1 ~ 4 4 0 0 0 4 を並列化しているので上記各統計を各パケット転送ユニット 4 4 0 0 ~ 4 4 0 7 (図 3)、4 4 1 0 ~ 4 4 1 3 (図 4) または各パケット検索ユニット 4 4 0 0 0 1 ~ 4 4 0 0 0 4 内で採取するようにした場合には、図 2 に示した RM 4 2 上で動作しているソフトウェア、或いは図 2 に示した RP 4 4 0 ~ 4 4 2 内にプロセッサを搭載し、このプロセッサ上でソフトウェアを動作させ、これらのソフトウェアが

各パケット転送ユニット4 4 0 0～4 4 0 7（図3）、4 4 1 0～4 4 1 3（図4）または各パケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4内に採取された統計情報を読み、並列化されている全パケット転送ユニット4 4 0 0～4 4 0 7（図3）、4 4 1 0～4 4 1 3（図4）または全パケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4に取られた統計情報を足し合わせるようにする。

【0 0 4 1】

図9に、本発明のパケット検索ユニット内で採取した統計情報をルーティングプロセッサ内が読み出し集計するという処理を表す概念図を示す。

図9では、パケット転送処理機構4 4 0 0 0内のパケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4は、それぞれ統計カウンタ4 4 0 0 0 1 3～4 4 0 0 0 4 3を備える。そして、パケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4内で採取した統計カウンタ4 4 0 0 0 1 3、4 4 0 0 0 2 3、4 4 0 0 0 3 3、4 4 0 0 0 4 3の値をR P 4 4 0内のプロセッサ4 4 5がレジスタバスを介して読み、集計する。図9ではパケットの収集に直接関係しない部分は省略してある。

【0 0 4 2】

また、統計機能ではないがパケットのデータ転送バイト数の測定が必要な場合として、ネットワーク業者がユーザからのパケットを受信する時にパケットのヘッダ条件に従いパケットを分類し、この分類(以後フローと呼ぶ)毎にデータ転送速度の監視を行い、このフローのデータ転送速度がこのフローを流しているユーザとの間で契約している帯域以上のパケットを廃棄したり輻輳時に優先して廃棄するようにする機能(帯域監視機能)、または、ネットワークユーザがネットワーク業者のネットワークにパケットを送信する時にルータから各フローに属するパケットデータを出力する速度をネットワーク業者との間で契約している帯域以内に制限する機能(シェーピング機能)がある。帯域監視機能またはシェーピング機能を実現するためには、ルータを流れるパケットの一定時間毎のデータ転送バイト数(データ転送速度)を測定する必要がある。

【0 0 4 3】

複数のパケット転送ユニット4 4 0 0～4 4 0 7（図3）、4 4 1 0～4 4 1 3（図4）または複数のパケット検索ユニット4 4 0 0 0 1～4 4 0 0 0 4を並列

化した場合には、一つのフローが複数のパケット転送ユニット4400～4407（図3）、4410～4413（図4）または複数のパケット検索ユニット440001～440004に分散されて処理されるので、一つのフロー全体のデータ転送速度を求めるためには、各パケット転送ユニット4400～4407（図3）、4410～4413（図4）または各パケット検索ユニット440001～440004を流れるフローのデータ転送速度を共通の個所に採取する必要がある。一つのフローのデータ転送速度を共通の個所に採取するために、各フローに番号を付け、各データ転送ユニット4400～4407（図3）、4410～4413（図4）または各データ検索ユニット440001～440004からパケットヘッダ検索結果の一つとしてフローの番号を出力し、転送パケット再配列ユニット4421、4423、4431、4433または検索パケット再配列ユニット440006内で、フローの番号毎にフローのデータ転送速度を測定するようにし、帯域監視機能実現時にはフローのデータ転送速度に従い、（例えば、閾値以上になったら）パケットを廃棄したり、パケットを廃棄されやすくしたりする。さらに、フロー番号毎にトラフィック統計を求めることもできる。また、シェーピング機能実現時にはフローのデータ転送速度がネットワーク業者との間の契約帯域を超えないように出力を制限するようにする。

【0044】

図10に、本発明の複数のパケット検索ユニットが検索したフロー番号に従い検索パケット再配列ユニットがフロー毎の転送速度を測定する部分を抜粋したパケット転送処理機構の内部構成図を示す。検索パケット再配列ユニット440006は、フロー番号に対応して、パケット数及びバイト数等のパケットデータ転送速度に関する情報を記憶したフロー毎の転送速度カウンタ4400060を備える。パケット検索ユニット440001～440004がパケット検索結果の一つとしてフロー番号を含む情報を出力し、検索パケット再配列ユニット440006がフローの番号毎の転送速度カウンタ4400060に各フローのパケットデータ転送速度を測定する。図10ではフロー毎のパケットデータ転送速度の測定に直接関係しない部分は省略してある。

【0045】

つぎに、フロー検索について説明する。図14に、フロー検索テーブルの説明図を示す。

このフロー検索テーブルは、一例として、図示のように、比較フィールド101として、転送元IPアドレス、宛先IPアドレス、パケット長、IP優先度、IP上位プロトコル、送達確認フラグ、転送先TCP/UDPポート、宛先TCP/UDPポートを含む。また、アクションフィールド102としては、フィルタ（通過／廃棄）、トンネル（カプセル化する／しない）、QoS（遅延クラス、廃棄クラス、帯域等）が記憶されている。

【0046】

つぎに、QoSフロー検索の具体的方法を説明する。ここでは、一例としてQoSフロー検索に注目して説明するが、各種フロー検索についても、同様に実施することができる。なお、アクションフィールド102には、これら各フローの制御情報が混在して記憶されていても良いし、また、各フロー毎にフロー検索テーブルが設けられるようにしても良い。

【0047】

まず、リニアサーチ方式について説明する。アクションのひとつとしてのパケットのQoS制御情報を判定する際、予め設定されたエントリをエントリテーブルの上から順に読みだす。そして、パケットのヘッダ部等の値とエントリの比較フィールド101内の有効なフロー条件に全て一致したか否かを判定する。なお、一致した場合にはエントリ内のアクションフィールド102内のQoS制御情報をパケットのQoS制御情報と判定し、QoSフロー検索を終了する。フロー条件に一致する場合は検索されると、アクションフィールド102内のQoS制御情報をQoS制御情報と決定し次のエントリの検索を実行せずフロー検索を終了する。

【0048】

以上のようなリニアサーチ方式では、大量のエントリが設定されるネットワークにおいては高速にQoS制御やフィルタリングを実行することが困難となる場合がある。そこで、大量のエントリが設定された場合でも、リニアサーチ方式と比較して高速にフロー検索を行うことができる、入力回線限定方式等を採用す

るとよい。以下、入力回線限定方式の概要を説明する。入力回線限定方式では、リニアサーチ方式の比較フィールドを構成する入力回線番号に一致するエントリだけを検索し、高速化を図るようにする。

【 0 0 4 9 】

図 1 5 に、第 1 の入力回線限定方式の説明図を示す。第 1 の入力回線限定方式においては、リニアサーチ方式の比較フィールドから入力回線番号と入力回線番号有効ビットを削除したエントリ 5 1 1 - i を入力回線毎に設定しておく。フロー条件部 5 2 1 - i は、例えば、送信元あるいは宛先ユーザを識別する条件は SIP と DIP の上限値と下限値である SIP 上限値 5 0 1、SIP 下限値 5 0 2、DIP 上限値 5 0 3、DIP 下限値 5 0 4 と、SIP と DIP の上限値と下限値が有効であることを示す IP 有効ビット 5 6 2 と、送信元ポートである SPORT 5 0 5 と、宛先ポートである DPORT 5 0 6 と、前記 SPORT 5 0 5 と DPORT 5 0 6 が有効であることを示すポート有効ビット 5 6 3 等を含む。Q o S 制御情報部 5 3 0 - i は、例えば、優先転送機能で使用する Q o S 制御情報である Q o S 制御情報 5 0 7 を含む。フロー条件である入力回線番号が一致したエントリ 5 1 1 - i だけ検索するためエントリ 5 1 1 - i 内に入力回線番号は必要ない。フロー検索時には、パケットが入力した入力回線に割り当てられたエントリ 5 1 1 - i のみ検索する。

【 0 0 5 0 】

ここで、上述の第 1 の入力回線限定方式では、入力回線番号に関係ないエントリ 5 1 1 - i を設定する（例えば「全ての入力回線から入力された T e l n e t のトラヒックは高優先」と設定する）場合、エントリ 5 1 1 - i を入力回線数（= N）設定する必要がある、エントリテーブルを実現するメモリの効率が悪くなる場合がある。そこで、以下に、一層高速な入力回線限定方式について説明する。

【 0 0 5 1 】

図 1 6 に、第 2 の入力回線限定方式の説明図を示す。第 2 の入力回線限定方式では、エントリテーブル 7 5 0 のアドレスであるリスト 5 4 0 を、入力回線毎にリストテーブル 7 6 0 に設定しておく。例えば、リストテーブルアドレスが " 1 " のリスト 5 4 0 - 1 1 はエントリ 5 1 1 - 1 のアドレスであり、リストテーブ

ルアドレスが” 2 ” のリスト 5 4 0 - 1 2 はエントリ 5 1 1 - H のアドレスである。フロー検索時には、パケットが入力された入力回線に割り当てられたリスト 5 4 0 だけ読みだし、このリスト 5 4 0 がポイントするエントリ 5 1 1 - i を読み出す。ビット幅の小さなリスト 5 4 0 (例えば、1 0 2 4 エントリ持った時にも 1 0 b i t 程度) を入力回線毎に所持し、ビット幅の大きなエントリ 5 1 1 - i を各入力回線で共有すれば、エントリテーブルを実現するメモリを有効に使用することができる。このため、高速化を実現しつつ、多数のエントリ 5 1 1 - i を設定することが可能となる。

【 0 0 5 2 】

フロー検出方式の他の実施例として出力回線限定方式がある。出力回線限定方式では、フロー条件である出力回線番号が一致するエントリ 5 1 1 - i のみを、上述の入力限定方式と同様に処理することで、フロー検出の高速化を実現する。また、フロー条件としてヘッダ情報内の入力回線番号の代わりに MAC (Media Access Control) レイヤーのソースアドレスである。SAMAC を用いた際の SAMAC 限定方式がある。SAMAC 限定方式では、SAMAC のグループである SAMAC グループを定義し、SAMAC グループの識別子である SAMAC 識別子でエントリを限定することで、上述の入力限定方式と同様にフロー検索を実行することができる。

【 0 0 5 3 】

なお、図 3、図 4 に示したパケット転送ユニットの並列化と図 7 に示したパケット検索ユニットの並列化を同時に実行することも可能であり、同時に実行した場合、各 R P 内のパケット検索ユニットの並列度は、パケット転送ユニットの並列度とパケット検索ユニットの並列度の掛け算になる。

【 0 0 5 4 】

【発明の効果】

本発明によると、パケットのヘッダ情報からパケットの宛先を検索しパケットを転送するという処理を実行するパケット転送ユニットを各 R P 内に複数個設け、N I F から入ってきたパケットおよび C S W から入ってきたパケットを、パケット単位にパケット転送ユニットに振り分けることにより、パケットの転送処理ユニットの並列度に応じパケット転送処理を高速化できる。

【 0 0 5 5 】

本発明によると、N I Fから入ってきたパケットをC S Wに転送する処理を行うパケット転送ユニットと、C S W経由で入ってきたパケットをN I Fに転送する処理を行うパケット転送ユニットを別々に設けることにより、それぞれのパケット転送ユニットが行う処理が限定され、パケット転送ユニットの構成を簡単化できる。本発明によると、各パケット転送ユニットにN I FからC S Wへの転送とC S WからN I Fへの転送のどちらの転送処理も行わせるようにすることにより、各パケット転送ユニットが処理をしていない時間を少なくしパケット転送ユニット当りのルータの処理性能を向上できる。

【 0 0 5 6 】

本発明によると、転送パケット振り分けユニットのパケット振り分け方法として、パケットが入ってきた順にパケット転送ユニットに順番に振り分ける方法(ラウンドロビン法)を採用することにより、パケットの振り分け方法が簡単であり実現が容易とすることができる。また、本発明によると、空いているパケット転送ユニットに振り分ける方法(アウトオブオーダ法)を採用することにより、パケット転送ユニットのパケット処理時間がパケット毎にばらついた場合でも、パケット転送ユニットのアイドル時間を短くし、ルータのトータル性能として見たときのパケット転送ユニット当りのパケット転送処理性能を良くすることができる。さらに、本発明によると、パケットのヘッダ条件から振り分けるパケット転送ユニットを一意に決定する方法(ハッシュ法)を採用することにより、パケットの転送元から転送先に連続して送られる一連のパケットが全て同じパケット転送ユニットに振り分けるので、パケットを送出する前にパケットの順番を再配列しなくてもパケットの追い抜きが発生しないようにすることができる。さらに、ハッシュ法において、各パケット転送ユニットにパケットができるだけ均等に振り分けられるように、ハッシュ関数としてパケットヘッダ条件から作成された擬似乱数を使用することにより、パケット転送ユニット間での処理パケット数の時間的ばらつきを抑えることができる。

【 0 0 5 7 】

本発明によると、パケット再配列ユニットは、各パケット転送ユニットが出力し

たパケットを R P へのパケットの入力順に再配列するので、パケットの追い抜きが発生しないようにできる。また、パケットの振り分け法がアウトオブオーダー法の場合に、パケットの振り分け時にパケットにシーケンス番号を付け、パケット転送ユニット内をパケットと共に持ち回ることにより、パケット再配列ユニットがパケットの入力順を知ることができ、パケットを入力順に再配列することができる。

【 0 0 5 8 】

本発明によると、各パケット転送ユニット内に転送パケット振り分けユニットからパケットを受けてからパケット転送処理を開始するまでの間待ったためのキュー、および、パケット転送処理ユニットが処理を終了してから転送パケット再配列ユニットにパケットを送出するまでの間待ったためのキューを設けることにより、各パケット転送ユニットの各パケットの処理時間の長短、および、転送パケット振り分けユニットの振り分け方がハッシュ法の場合に各パケット転送ユニットにパケットを振り分けるときの時間的偏りを吸収し、各パケット転送ユニットが処理をしていない時間を少なくしパケット転送ユニット当りのルータの処理性能を向上できる。

【 0 0 5 9 】

本発明によると、パケット転送ユニットでなく、パケット転送ユニットが行う処理の内、パケットヘッダの検索処理という限定された処理だけを行うパケット検索ユニットのみを並列化することにより、ハードウェアの増加を抑えつつ、パケット毎に必要な処理の内複雑でボトルネックになり易い処理であるパケットヘッダ検索処理を高速化できる。

【 0 0 6 0 】

本発明によると、パケット転送ユニットの並列化とパケット検索ユニットの並列化を同時に実行することにより、ルータ内でパケット毎に必要な処理の並列度を、パケット転送ユニットの並列度とパケット検索ユニットの並列度まで高めることができ、ルータのトータル性能をさらに上げることができる。

【 0 0 6 1 】

本発明によると、複数のパケット転送ユニットまたは複数のパケット検索ユニッ

トのそれぞれに統計情報を採取し、ルータ全体を管理するソフトウェアが各パケット転送ユニットまたはパケット検索ユニット内の統計情報を読み、ソフトウェア処理により集計するか、あるいは、各 R P 毎にプロセッサを設け、このプロセッサ上で動作するソフトウェアが各パケット転送ユニットまたはパケット検索ユニット内の統計情報を読み、ソフトウェア処理により集計することにより、ルータ全体としての統計情報を得ることができる。

【 0 0 6 2 】

本発明によると、複数のパケット転送ユニットまたは複数のパケット検索ユニットのそれぞれが、パケットのヘッダ条件毎にパケットを分類し、この分類(以後フローと呼ぶ)に番号を付け、各データ転送ユニットまたは各データ検索ユニットからフローの番号を出力し、転送パケット再配列ユニットまたは検索パケット再配列ユニット内で、フローの番号毎にフローのデータ転送速度を測定することにより、全パケット転送ユニット、あるいは全パケット検索ユニットで処理したトータルとしてのフロー毎のデータ転送速度を測定でき、データ転送速度に従いパケットを廃棄したり輻輳時にパケットを廃棄され易くすること、あるいはパケットの送出時にパケットの送出速度を制限することができる。

【図面の簡単な説明】

【図 1】

本発明のルータが適用されるネットワークシステム全体図。

【図 2】

本発明のルータ全体の内部構成ブロック図。

【図 3】

パケット転送ユニットを並列化した本発明のルーティングプロセッサ内部ブロック図(転送方向によりパケット転送ユニットを独立に持つ場合のブロック図)。

【図 4】

パケット転送ユニットを並列化した本発明のルーティングプロセッサ内部ブロック図(一つのパケット転送ユニットが、両方向のパケット転送処理を行う場合のブロック図)。

【図 5】

本発明の packets 転送ユニットの内部ブロック図。

【図 6】

本発明の packets 検索ユニットを並列化しない場合の packets 転送ユニットとその内部の packets 転送処理機構の内部ブロック図。

【図 7】

本発明の packets 検索ユニットを並列化した場合の packets 転送ユニットとその内部の packets 転送処理機構の内部ブロック図。

【図 8】

本発明の packets 検索ユニットの内部ブロック図。

【図 9】

本発明の packets 検索ユニット内で採取した統計情報をルーティングプロセッサ内が読み出し集計するという処理を表す概念図。

【図 1 0】

本発明の複数の packets 検索ユニットが検索したフロー番号に従い検索 packets 再配列ユニットがフロー毎の転送速度を測定する部分を抜粋した packets 転送処理機構の内部ブロック図。

【図 1 1】

本発明の転送 packets 振り分けユニットの振り分け方法がラウンドロビン法の場合の packets 転送ユニットの packets 転送処理を表すタイムチャート。

【図 1 2】

本発明の転送 packets 振り分けユニットの振り分け方法がアウトオブオーダー法の場合の packets 転送ユニットの packets 転送処理を表すタイムチャート。

【図 1 3】

本発明の転送 packets 振り分けユニットの振り分け方法がハッシュ法の場合の packets 転送ユニットの packets 転送処理を表すタイムチャート。

【図 1 4】

フロー検索テーブルの説明図。

【図 1 5】

第 1 の入力回線限定方式の説明図。

【図 1 6】

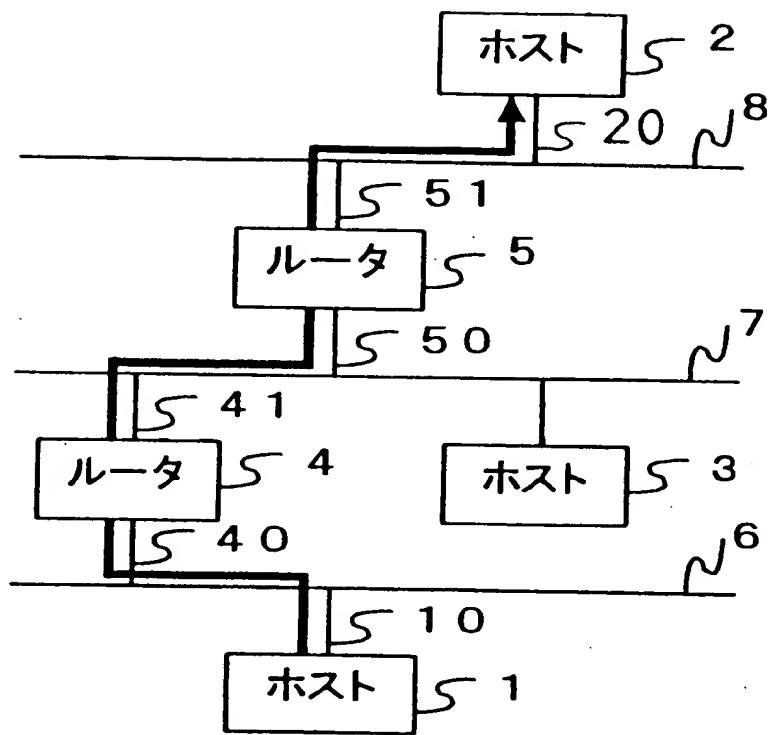
第 2 の入力回線限定方式の説明図。

【符号の説明】

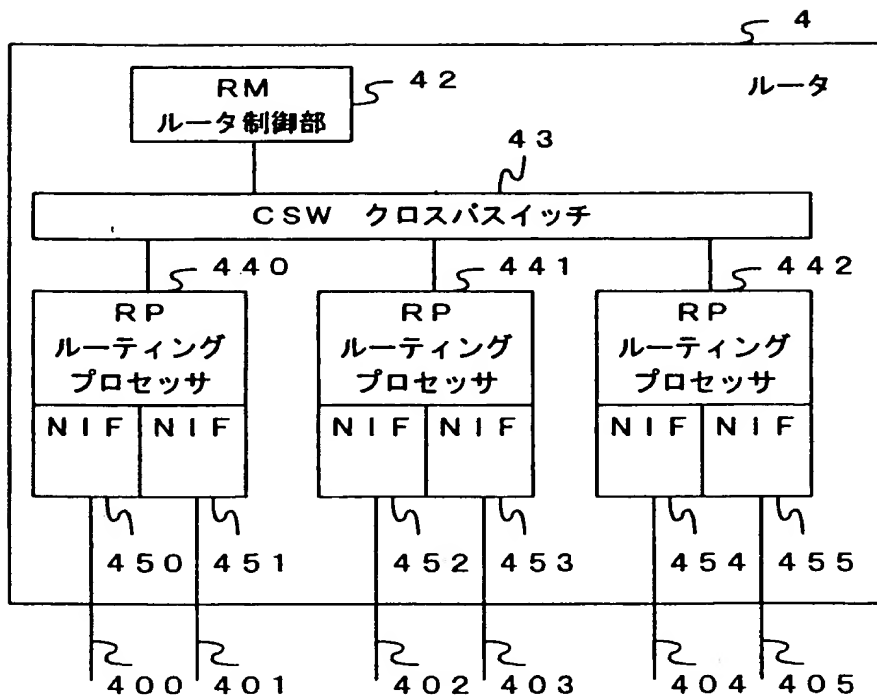
1 ～ 3 ホスト、 4 ～ 5 ルータ、 6 ～ 8 サブネットワーク、
4 0 0 ～ 4 0 5 ネットワークインターフェース、 4 2 ルータ制御部 (RM)、
4 3 クロスバスイッチ (CSW)、 4 4 0 ～ 4 4 2 ルーティングプロセッサ、
4 5 0 ～ 4 5 5 ネットワークインターフェースボード (NIF)、
4 4 0 0 ～ 4 4 0 7 パケット転送ユニット、 4 4 2 0、 4 4 2 2 転送パケット
ト振り分けユニット、 4 4 2 1、 4 4 2 3 転送パケット再配列ユニット、
4 4 1 0 ～ 4 4 1 3 パケット転送ユニット、 4 4 3 0、 4 4 3 2 転送パケット
ト振り分けユニット、 4 4 3 1、 4 4 3 3 転送パケット再配列ユニット、
4 4 0 0 0 パケット転送処理機構、 4 4 0 0 1 転送パケット処理待ちキュー
、 4 4 0 0 2 転送パケット再配列処理待ちキュー、
4 4 0 0 0 0 パケットヘッダ操作機構、 4 4 0 0 0 1 ～ 4 4 0 0 0 4 パケッ
ト検索ユニット、 4 4 0 0 0 5 検索パケット振り分けユニット、 4 4 0 0 0 6
 検索パケット再配列ユニット、
4 4 0 0 0 1 0 パケット検索処理機構、 4 4 0 0 0 1 1 検索パケット処理待
ちキュー、 4 4 0 0 0 1 2 検索パケット再配列処理待ちキュー、
4 4 0 0 0 1 3、 4 4 0 0 0 2 3、 4 4 0 0 0 3 3、 4 4 0 0 0 4 3 統計カウ
ンタ、 4 4 5 プロセッサ、
4 4 0 0 0 6 0 フロー毎の転送速度カウンタ

【書類名】図面

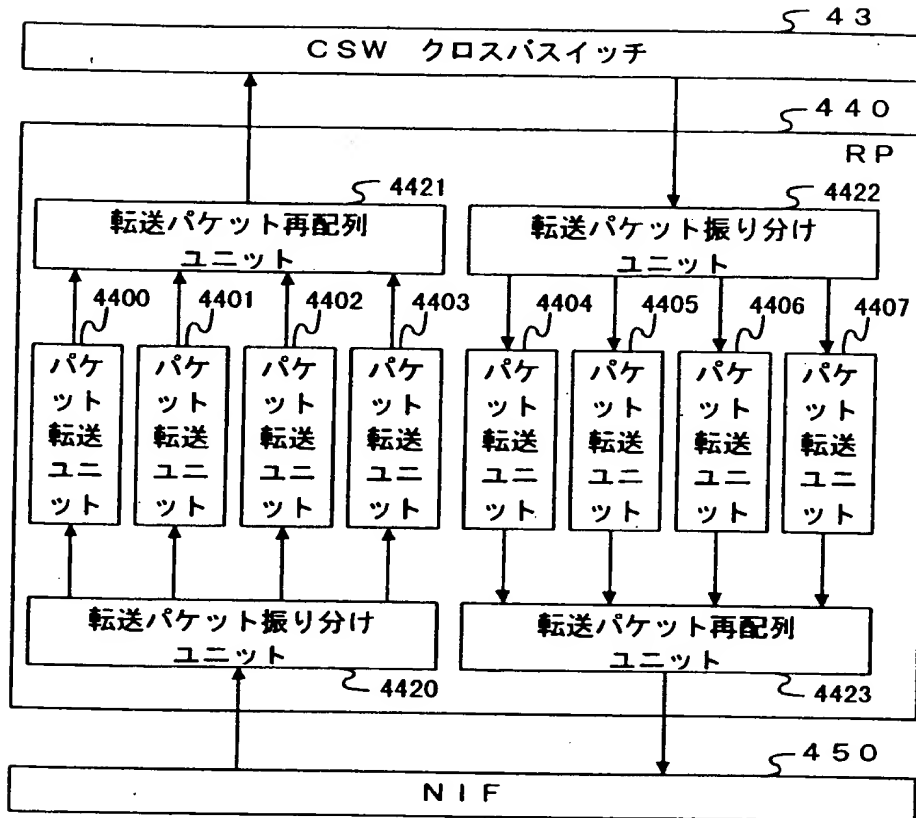
【図1】



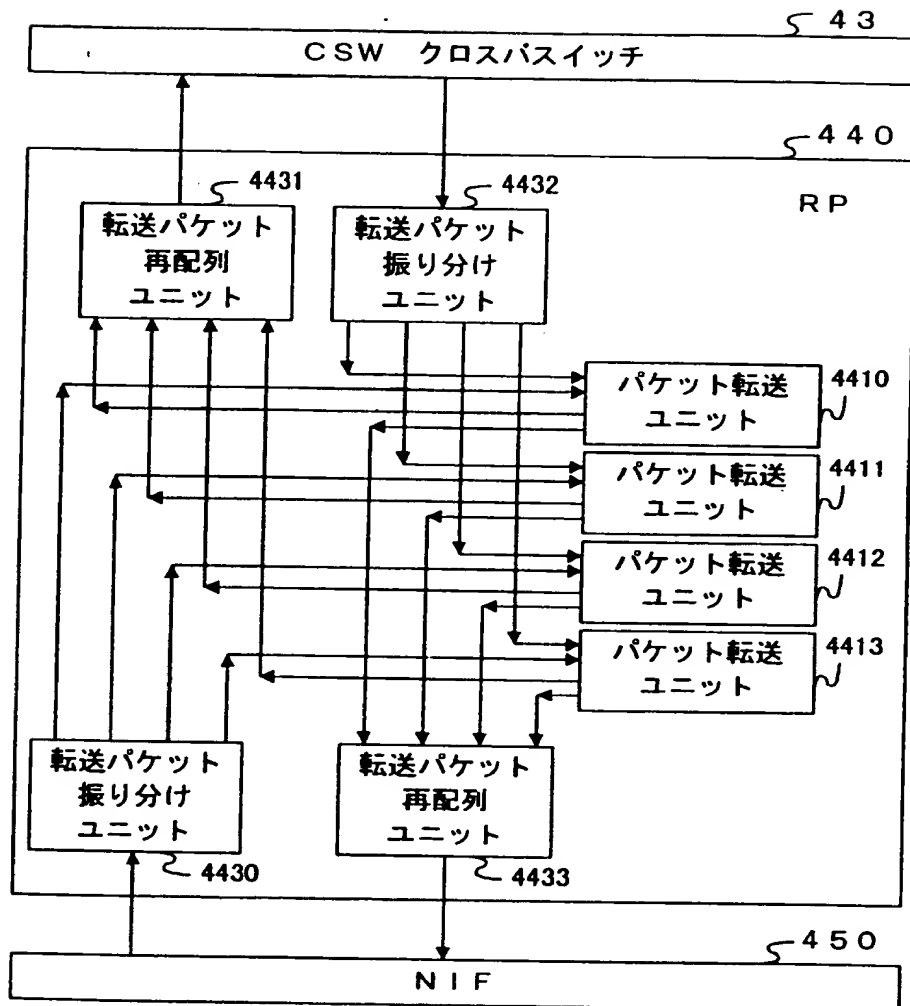
【図 2】



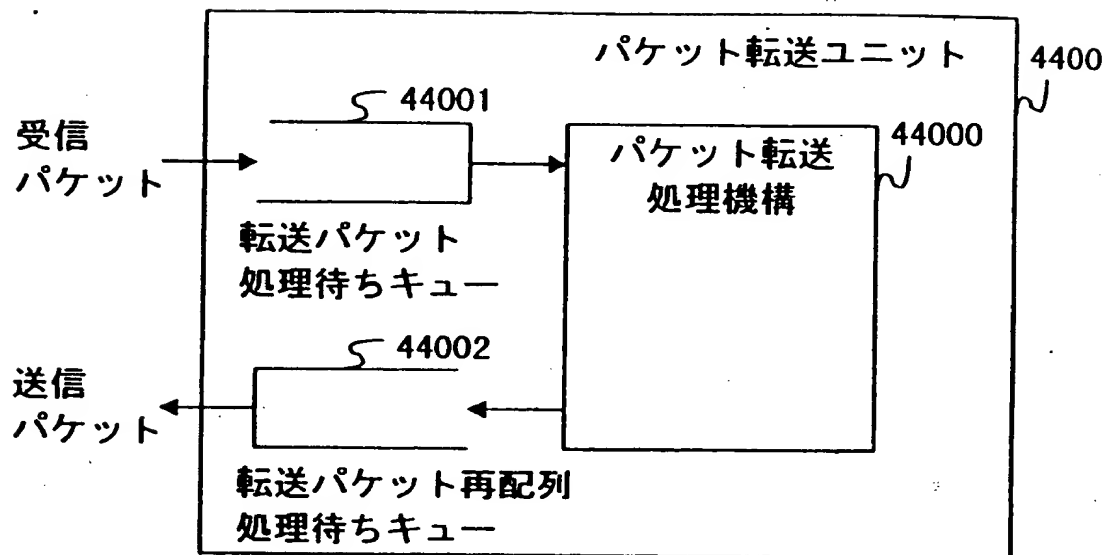
【図 3】



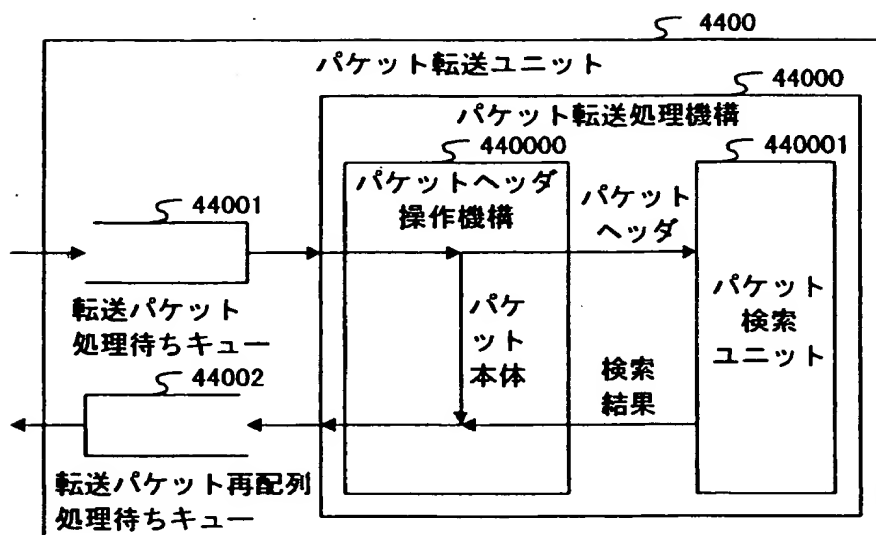
【図 4】



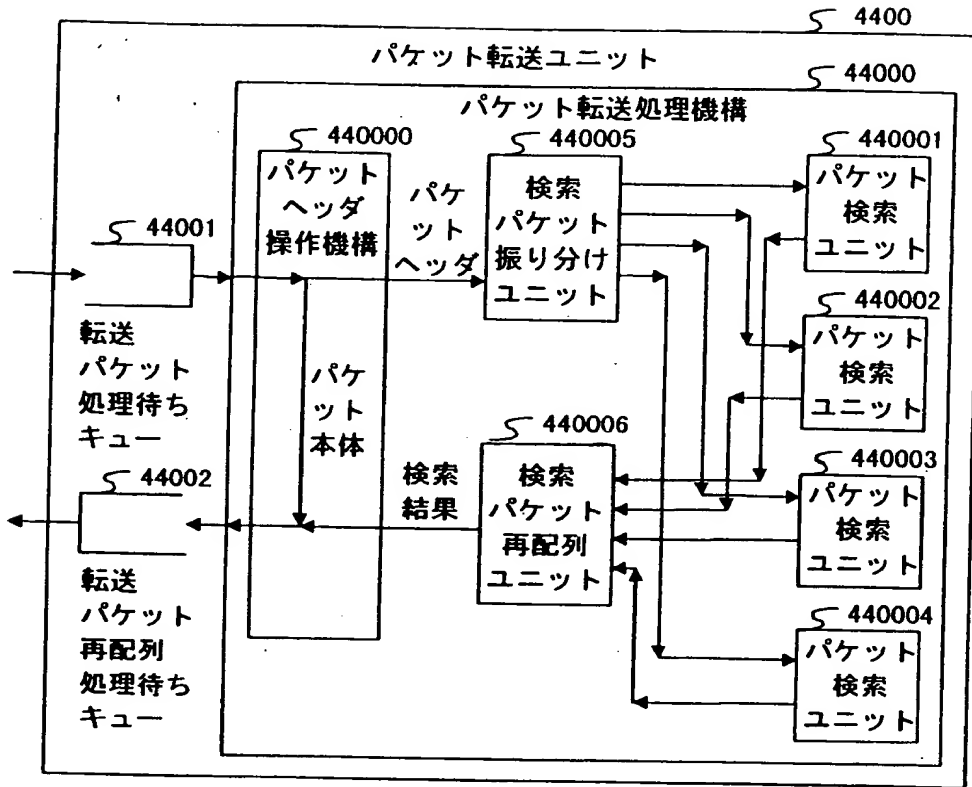
【図 5】



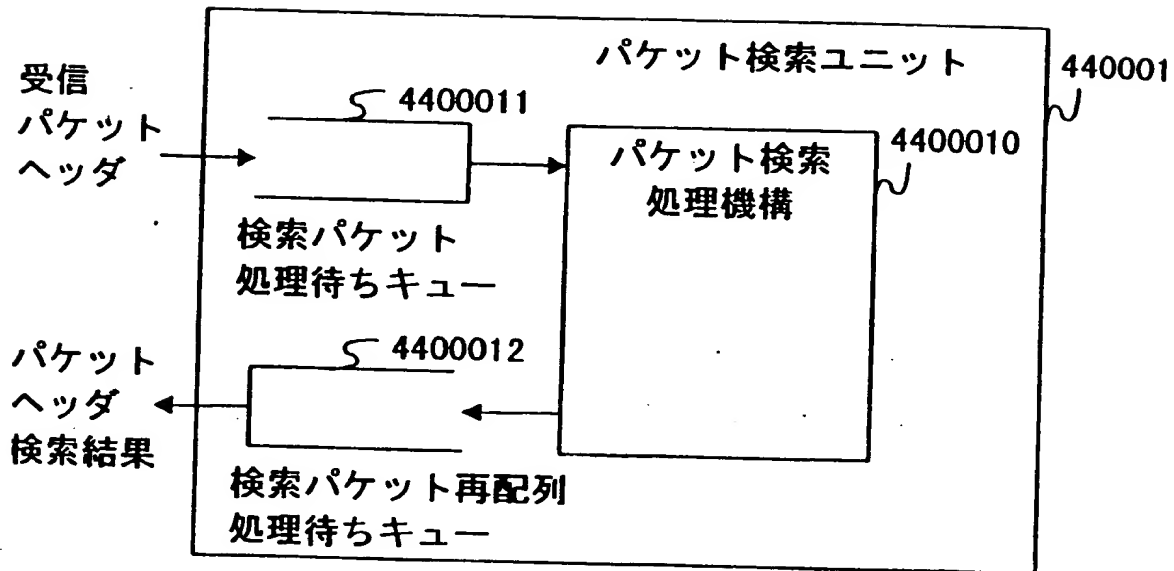
【図 6】



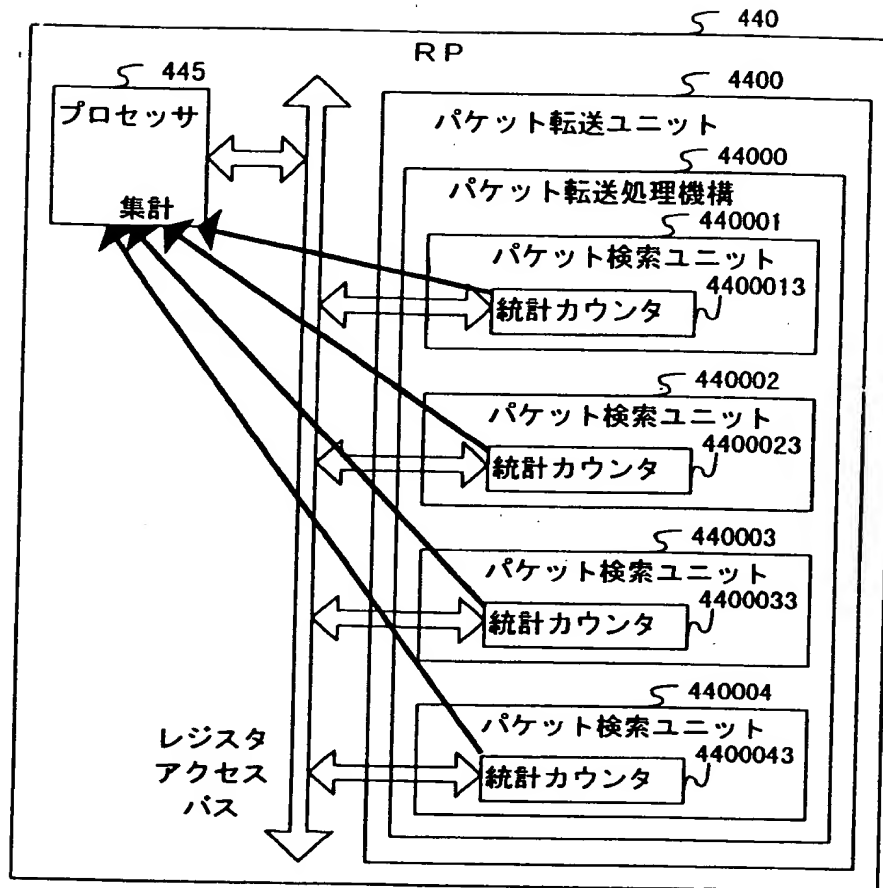
【図 7】



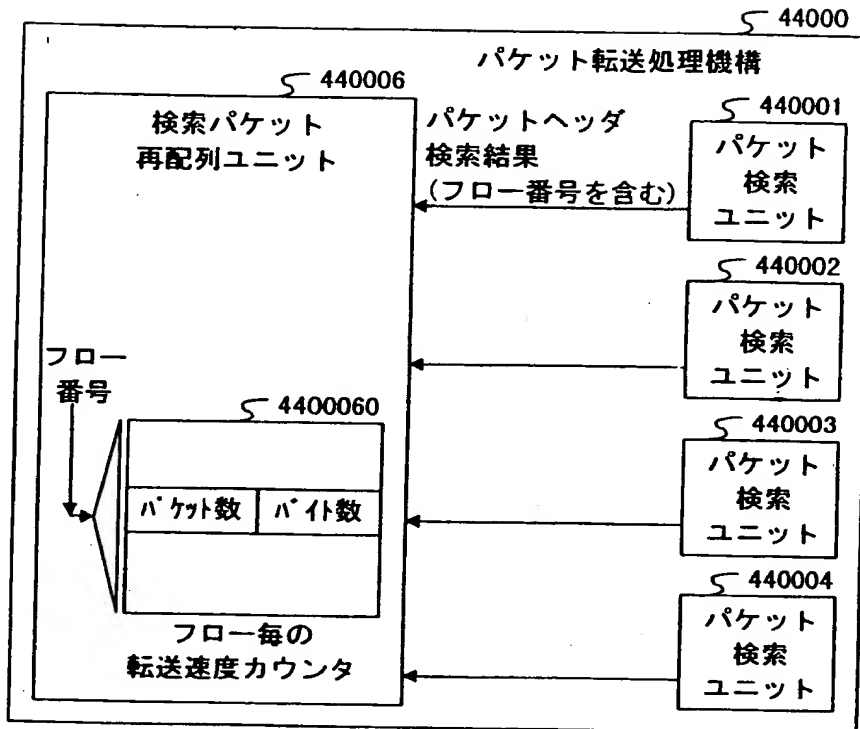
【図 8】



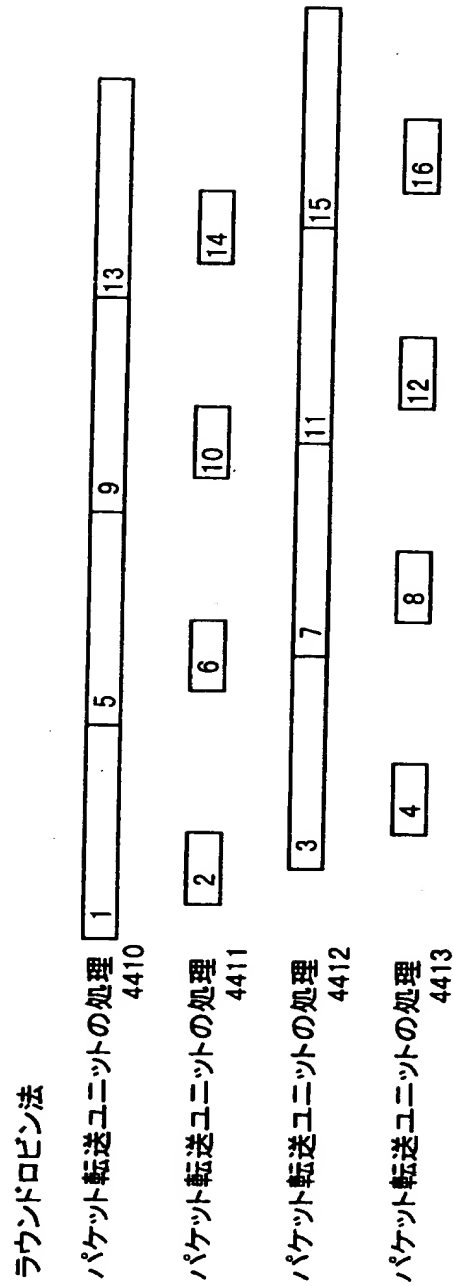
【図 9】



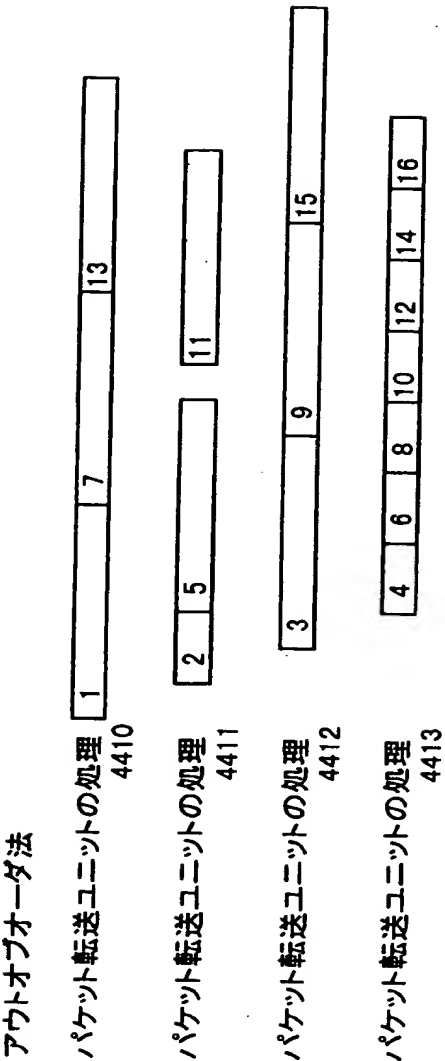
【図 1 0】



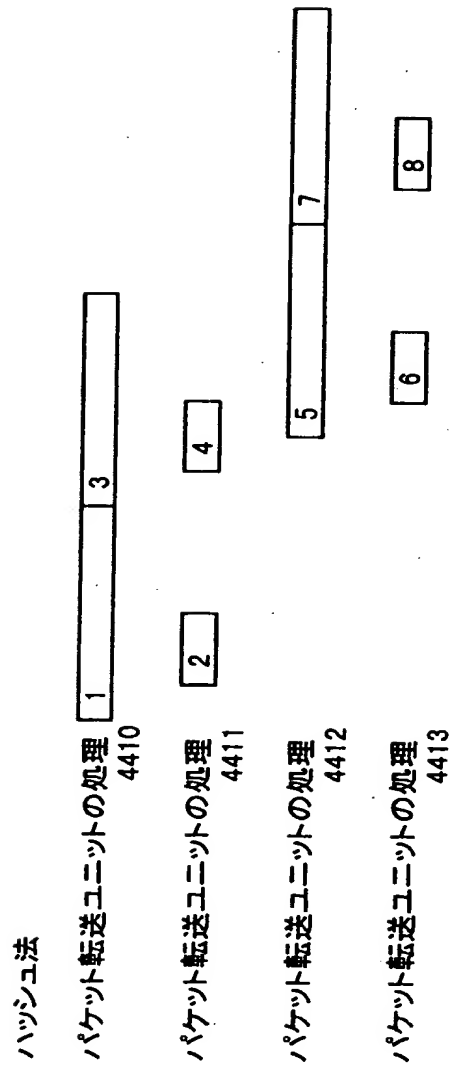
【図 1 1】



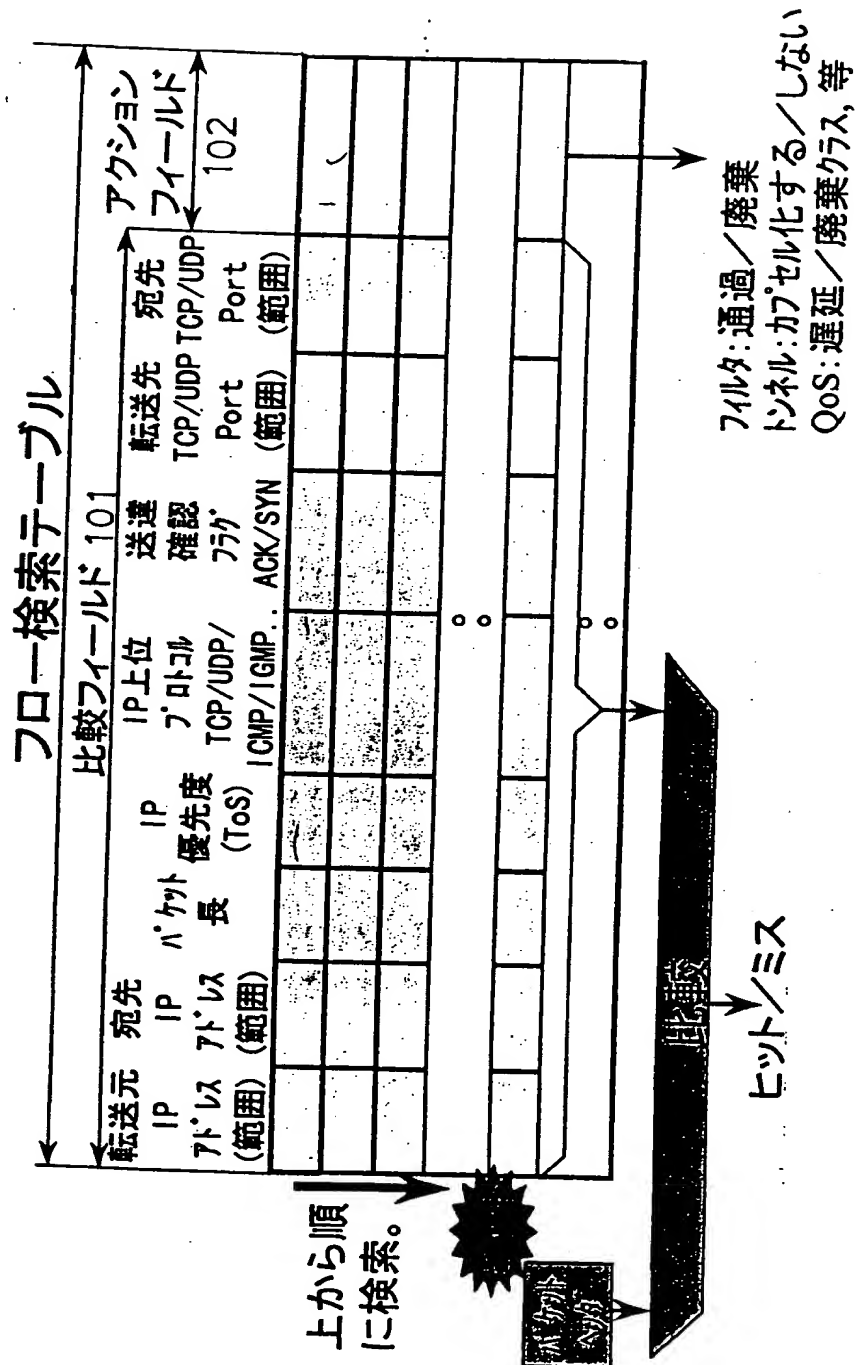
【図 1 2】



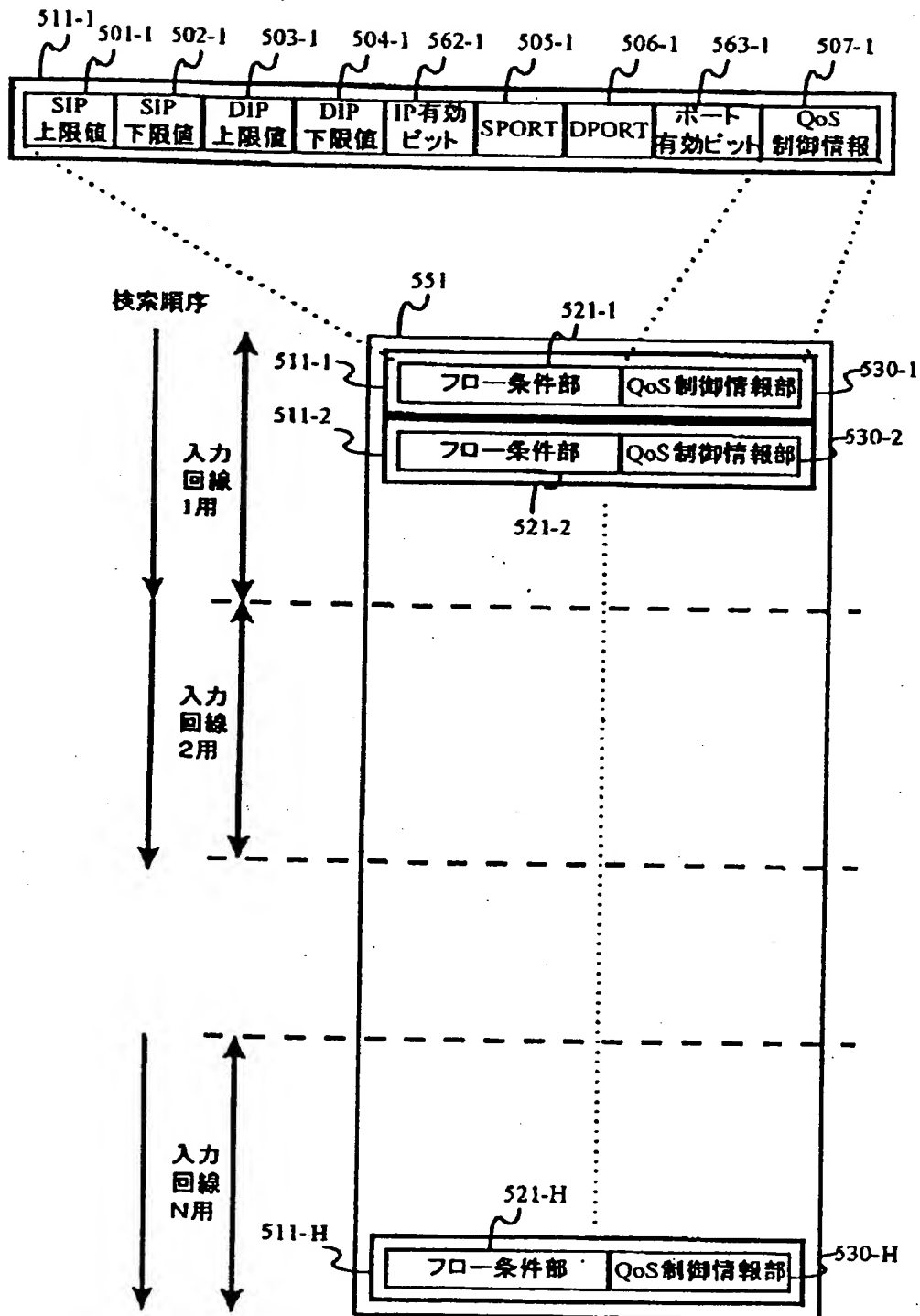
【図 1 3】



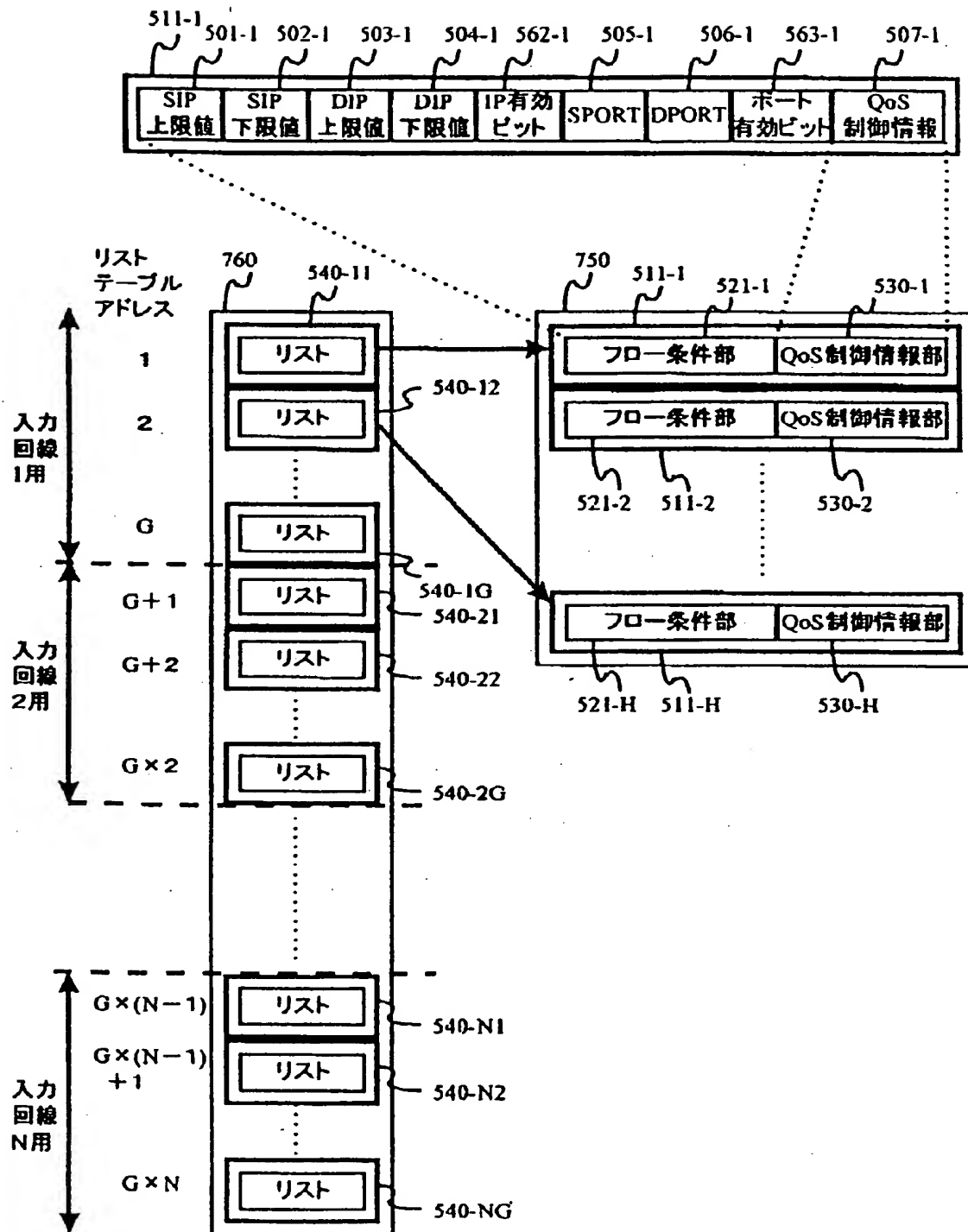
【図 14】



【図 1 5】



【図 1 6】



【書類名】 要約書

【要約】

【課題】 パケットの転送処理機構、検索処理機構を並列化し、ハードウェア要素の高速化以上にネットワーク中継装置全体の中継処理性能を向上させる。

【解決手段】 ルータ内でパケットの転送処理を行うパケット転送ユニット 4 4 0 0 ～ 4 4 0 7 を並列化し、並列化した各パケット転送ユニット 4 4 0 0 ～ 4 4 0 7 へパケットを振り分けるパケット振り分けユニット 4 4 2 0、4 4 2 2 と、パケット転送ユニット 4 4 0 0 ～ 4 4 0 7 の出力を再配列するパケット再配列ユニット 4 4 2 1、4 4 2 3 を設ける。さらに、パケット転送ユニット 4 4 0 0 ～ 4 4 0 7 内でパケットヘッダを検索するパケット検索ユニットを並列化する。

【選択図】 図 3

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所

出 願 人 履 歴 情 報

識別番号 [0 0 0 1 5 3 4 5 4]

1. 変更年月日 1 9 9 9 年 8 月 2 5 日

[変更理由] 住所変更

住 所 神奈川県足柄上郡中井町境 4 5 6 番地

氏 名 株式会社日立インフォメーションテクノロジー